JOURNAL OF NUMBER THEORY 10, 250–262 (1978)

An Almost-Prime Sieve

DOUGLAS HENSLEY*

Department of Mathematics, Texas A & M University, College Station, Texas 77843

Communicated by S. Chowla

Received October 11, 1977

We compare and contrast three methods for estimating the number of integers in an interval of length x which have fewer than k distinct prime factors less than z, with special attention to the case $k = 2$. An iterative method based on the case $k = 1$ is simplest. If z is sufficiently small compared to x one may use a kind of Brun sieve. Selberg's sieve method gives a good estimate for $k = 2$ but leads into technical difficulties as k increases.

1. INTRODUCTION

Our main interest is to see how certain classical sieve methods perform in a setting where the usual hypotheses are absent or modified. For the sake of simplicity we do not push our ideas to their limits.

Let $\mathscr{M}$ be the set of integers in the interval $(y, y + x]$, let $\mathscr{P}$ be the set of primes less than z, and let $k \geqslant 2$ be an integer. Let $\mathscr{N}$ be the set of all members of $\mathscr{M}$ with fewer than k distinct prime factors in $\mathscr{P}$. We describe here briefly three approaches to bounding $\#\mathscr{N}$, the number of elements of $\mathscr{N}$, in terms of x, z, and k. Each approach is discussed in more detail in a subsequent section.

The iterative method assumes that one has a function $E_{k-1}(x, z)$ such that $\#\{n \in (y, y + x]: n$ has fewer than $k - 1$ distinct prime factors less than $z\} \leqslant E_{k-1}(x, z)$ for all x, z. We show in Section 1 that one may then take

$$E_k(x, z) = E_{k-1}(x, z) + \frac{1}{k-1} \sum_{p<z} E_{k-1}\left(\frac{x}{p}, z\right)$$

as an upper bound for $\#\mathscr{N}$.

The combinatorial sieve proceeds from an identity and an inequality. Let $\mu(d)$ be the Möbius function and $\nu(d)$ the number of distinct prime factors of d. Let

$$\mu_k(d) = \mu(d)(-1)^{k-1} \binom{\nu(d) - 1}{k - 1}.$$

* Supported in part by NSF Grant MCS72-05055 AO4.

0022-314X/78/0102-0250$02.00/0

Then

$$\sum_{d|n} \mu_k(d) = \begin{cases} 1 & \text{if } \nu(d) < k \\ 0 & \text{if } \nu(d) \geqslant k. \end{cases}$$

Thus

$$\sum \frac{\mu_k(n)}{n^s} \zeta(s) = \sum_{\nu(n)<k} n^{-s}.$$

Further, if the above sum is truncated and taken only over d with $\nu(d) \leqslant m$ it alternates about the final value as m increases to $\nu(n)$. These facts with $k = 1$ were the basis of Brun's earliest sieve.

In the Selberg method one constructs a function $s^+(n)$ such that $s^+(n) = 1$ if $\nu(n) < k$, $s^+(n) \geqslant 0$ if $\nu(n) \geqslant k$. Then $\sum_{n\in(y,y+x]} s^+(n) \geqslant \#\mathcal{N}$.

2. Iteration

THEOREM 2.1. *Suppose for $k < K$, $E_k(x, z) \geqslant \#\mathcal{N}$ for any y. Then we may take*

$$E_K(x, z) = E_{K-1}(x, z) + \frac{1}{K-1} \sum_{p<z} E_{K-1}\left(\frac{x}{p}, z\right)$$

as an upper bound for $\#\{n \in (y, y+x]$: n has fewer than K distinct prime factors less than $z\}$.

Proof. Fix x, y, and z. Let $\pi = \prod_{p<z} p$. Let $\mathcal{N} = \{n \in \mathcal{N}\colon \nu((n, \pi)) < K\}$. Let $\mathcal{N}_p = \{n \in \mathcal{N}\colon p \mid n,\ \nu((n, \pi)) = K-1\}$. Then

$$\begin{aligned} \#\mathcal{N} &= \frac{1}{K-1} \sum_{p<z} \#\mathcal{N}_p + \#\{n \in \mathcal{N}\colon \nu((n, \pi)) < K-1\} \\ &\leqslant \frac{1}{K-1} \sum_{p<z} E_{K-1}(x/p, z) + E_{K-1}(x, z). \end{aligned}$$

Montgomery and Vaughan [3] give $E_1(x, x) = 2x/\log x$, and it is known that one cannot take $E_1(x, x)$ as small as $x/\log x$. On taking $E_1(x, x) = 2x/\log x$ in Theorem 2.1, we can take

$$E_k(x, x) = \frac{2x(\log\log x)^{k-1}}{(k-1)!\log x} (1 + \epsilon_k(x))$$

where for each fixed k,

$$\lim_{x\to\infty} \epsilon_k(x) = 0.$$

Remark. While we wish to bound $\max_y \sum_p \#\mathcal{N}_p$ in Theorem 2.1, the proof actually bounds the larger $\sum_p \max_y \#\mathcal{N}_p$. Yet our $E_k(x, x)$ is not larger than best possible by a factor of more than 2, as with $y = 0$ one has

$$\#\mathcal{N} \sim \frac{x(\log\log x)^{k-1}}{(k-1)!\log x}.$$

3. Combinatorics

In this section we give some exact formulas which can be made to play a part in Brun-type or Selberg-type sieves for almost primes. Let $x, y, z > 0$ and $k \in Z$. Let $\pi = \prod_{p<z} p$. Let $\mathcal{N}$ now be an arbitrary subset of $(y, y + x]$. We say $\mathcal{N}$ is *sifted* (k, z) (or when context permits, *sifted*) if there exists y' such that for $n \in \mathcal{N}$, $\nu((n + y', \pi)) < k$.

THEOREM 3.1. *The following are equivalent*:

(1) *$\mathcal{N}$ is sifted (k, z).*

(2) *There exist congruence classes $y_p \bmod p$ for $p < z$ such that for all $n \in \mathcal{N}$, $\#\{p < z: n \equiv y_p (\bmod p\} < k$.*

(3) *There exist infinitely many y' such that for all $n \in \mathcal{N}$, $\nu((n + y', \pi)) < k$.*

Proof. It is obvious that (3) implies (1). Conversely, if y' works in (1) then for $m \in Z$, $y' + m\pi$ does too, so (1) implies (3). If (1), (2) follows on taking $y_p \equiv -y \pmod p$ for $p < z$. Finally if (2), (1) follows via the Chinese remainder theorem with y' such that $y' \equiv -y_p \pmod p$ for $p < z$.

We shall need the following theorem due to Tijdeman, which can be derived from Feldman's theorem [1]:

THEOREM. *Let $\mathcal{P}$ be a set of primes and let $A(\mathcal{P}) = \{a: p \mid a \Rightarrow p \in \mathcal{P}\}$. Then there exists $C > 0$ such that if $a_1 < a_2 \in A(\mathcal{P})$ then $a_2 - a_1 > a_1/(\log a_1)^C$. (Actually we only need to know that $a_2 - a_1 \to \infty$ as $a_1 \to \infty$; this was known much earlier.)*

The following conjecture (A) illustrates the connection between sifted sets and almost primes.

Conjecture A (The linear case of Schinzel's Hypothesis H). If $(b_1 x + \beta_1)$, $(b_2 x + \beta_2), \ldots, (b_N x + \beta_N)$ are polynomials in x such that the b's and β's are integers and such that no prime divides $\prod_{i=1}^{n} (b_i x + \beta_i)$ for all integers x, then there are infinitely many x for which $(b_i x + \beta_i)$ is prime for $1 \leqslant i \leqslant N$.

THEOREM 3.2. *For fixed z and $\mathcal{N}$ if there are infinitely many y such that $\nu(y+n) \leqslant k$ for all $n \in \mathcal{N}$ then $\mathcal{N}$ is sifted (k, z).*

Proof. Take $\mathscr{P} = \{p \text{ prime}: p < z\}$. For sufficiently large y, no more than one $y + n \in A(\mathscr{P})$ for $n \in \mathcal{N}$. Thus for all save one $n \in \mathcal{N}$, $\nu((y+n, \pi)) < k$. Now if the distance between members of $A(\mathscr{P}) > x + y$ exceeds $\pi + 2x$, either $\nu((y+n,\pi)) < k$ for $n \in \mathcal{N}$ or $\nu(y + \pi + n, \pi) < k$ for $n \in \mathcal{N}$.

THEOREM 3.3. *If $\mathcal{N}$ is sifted (k, x) then on conjecture A there exist infinitely many y such that $\nu(y+n) \leqslant k$ for $n \in \mathcal{N}$.*

Proof. There exists y_1 such that $\nu((y_1 + n, \pi)) < k$ for $n \in \mathcal{N}$. Consider $\{y_1 + m\pi: m \text{ integer}\}$. Let $M_1 = \max_{p<z} \max_{n\in\mathcal{N}}$ (# times p divides $y_1 + n$). Let $M_2 = \prod_{p<z} p^{M_1+1}$. Then for $p < z$, $n \in \mathcal{N}$, m integer, the power of p which divides $y_1 + n =$ that power which divides $y_1 + n + mM_2$. Now for $n \in \mathcal{N}$, let $a_n = (y_1 + n, M_2)$. Then $a_n \in A(\mathscr{P})$, $a_n \mid (y_1 + n + mM_2)$ for $n \in \mathcal{N}$, m integer, and $((y_1 + n + mM_2)/a_n, \pi) = 1$ for $n \in \mathcal{N}$, m integer. Let $b_n = M_2/a_n$, $\beta_n = (y+n)/a_n$. Then for all m, $(b_n m + \beta_n, \pi) = 1$. If on the other hand $p > x \geqslant N$ then since p divides no b_n, there are m such that for $n \in \mathcal{N}$, $p \nmid (b_n m + \beta_n)$. Thus the conditions for the conjecture (A) are satisfied and on (A), there exist y such that $\nu(y+n) \leqslant k$ for $n \in \mathcal{N}$.

Now recall from the Introduction that $\mu_k(d) = \mu(d)(-1)^{k-1}\binom{\nu(d)-1}{k-1}$.

THEOREM 3.4. *For $k \geqslant 1$,*

$$\mu_k(d) = \sum_{i=0}^{k-1} \mu(d)(-1)^i \binom{\nu(d)}{i}.$$

Proof. If $\nu(d) < k$ both sides are zero. If $\nu(d) \geqslant k$ the above sum can be written as $\sum_{i=0}^{k-1} \mu(d)(-1)^i(\binom{\nu(d)-1}{i-1} + \binom{\nu(d)-1}{i})$ which telescopes to $\mu(d)(-1)^{k-1}\binom{\nu(d)-1}{k-1} = \mu_k(d)$.

THEOREM 3.5. *If $\nu(n) < k$, $\sum_{d|n} \mu_k(d) = 1$; if $\nu(n) \geqslant k$, $\sum_{d|n} \mu_k(d) = 0$.*

Proof. If $\nu(n) < k$ then $\mu_k(1) = 1$. If, moreover, $1 < d \mid n$ then

$$\mu_k(d) = \sum_{i=0}^{k-1} \mu(d)(-1)^i \binom{\nu(d)}{i} = \sum_{i=0}^{\nu(d)} \mu(d)(-1)^i \binom{\nu(d)}{i} = (1-1)^{\nu(d)} = 0.$$

Thus if $\nu(n) < k$, $\sum_{d|n} \mu_k(d) = 1$.

Conversely if $\nu(n) \geqslant k$ then

$$\sum_{d \mid n} \mu_k(d) = \sum_{m=0}^{\nu(n)} \sum_{i=0}^{k-1} (-1)^{i+m} \binom{m}{i} \binom{\nu(n)}{m}$$

$$= \sum_{i=0}^{k-1} (-1)^i \sum_{m=0}^{\nu(n)} (-1)^m \binom{\nu(n)}{i} \binom{\nu(n)-i}{m-i}$$

$$= \sum_{i=0}^{k-1} (-1)^i \binom{\nu(n)}{i} \sum_{m=i}^{\nu(n)} (-1)^m \binom{\nu(n)-i}{m-i} = 0,$$

as the inner sum is $=(1-1)^{\nu(n)-i} = 0$.

We now consider the effect of truncating the sum above. Let $\theta_k(n, m) = \sum_{d \mid n, \nu(d) \leqslant m} \mu_k(d)$.

THEOREM 3.6. *The following hold for* $\theta_k(n, m)$:

(1) *If* $\nu(n) < k$ *then* $\theta_k(n, m) = 1$.

(2) *If* $m < k \leqslant \nu(n)$ *then* $\theta_k(n, m) = 1$.

(3) *If* $k \leqslant \nu(n) \leqslant m$ *then* $\theta_k(n, m) = 0$.

(4) *If* $k \leqslant m < \nu(n)$ *and* $k + m$ *is odd,* $\theta_k(n, m) \geqslant 0$ *while if* $k + m$ *is even,* $\theta_k(n, m) \leqslant 0$.

The proof of the theorem rests on the following.

LEMMA. *If* $k \leqslant m < \nu(n)$, P *is the product of* k *distinct prime divisors of* n, $d \mid n/P$ *and* $C(d) = \{d' : d' = dq$, *for some* $q \mid P\}$ *then if* $k + m$ *is odd,* $\sum_{d' \in C(d), \nu(d') \leqslant m} \mu_k(d') \geqslant 0$ *while if* $k + m$ *is even, the sum above is* $\leqslant 0$.

To prove the lemma, let $\nu(d) = D$, $m - \nu(d) = E \geqslant 0$. Then

$$\sum_{\substack{d' \in C(d) \\ \nu(d') \leqslant m}} \mu_k(d) = (-1)^{k+D-1} \sum_{j=0}^{E} (-1)^j \binom{k}{j} \binom{D+j-1}{k-1}$$

$$= (-1)^{k+D+E-1} \binom{k+D-1}{E}$$

which proves the lemma since $D + E = m$.

We now prove Theorem 3.6. Both (1) and (2) follow from the definition of μ_k. Theorem 3.5 gives (3) so only (4) remains. If we write

$$\theta_k(n, m) = \sum_{\substack{d \mid n/P \\ \nu(d) \leqslant m}} \sum_{\substack{d' \in C(d) \\ \nu(d') \leqslant m}} \mu_k(d')$$

and apply the lemma to the inner sum, (4) follows.

Remark. This inequality (4) with $k = 1$ is the starting point for the simplest form of the Brun sieve.

THEOREM 3.7. *Let $e(x)$ denote $e^{2\pi i x}$. Then for d square free,*

$$\sum_{\substack{a=1 \\ \nu((a,d))\leqslant k}}^{d} e\left(\frac{a}{d}\right) = \mu_k(d).$$

Proof. With $k = 1$ this is the Hardy–Ramanujan identity. Suppose the theorem holds if $k \leqslant K$. We show that

$$\sum_{\{a:\nu((a,d))=K,1\leqslant a\leqslant d\}} e(a/d) = (-1)^k \mu(d) \binom{\nu(d)}{k}.$$

In view of Theorem 3.4 this will prove (3.7) by induction. Now

$$\sum_{\{a:\nu((a,d))=K,1\leqslant a\leqslant d\}} e\left(\frac{a}{d}\right) = K^{-1} \sum_{p\mid d} \sum_{\{a:1\leqslant a\leqslant d/p,\nu((a,d))=K-1\}} e\left(\frac{ap}{d}\right)$$

$$= K^{-1} \sum_{p\mid d} (-1)^{K-1} \mu(d/p) \binom{\nu(d)-1}{K-1}$$

by the induction hypothesis, and this is

$$= K^{-1}(-1)^K \mu(d)\, \nu(d) \binom{\nu(d)-1}{K-1} = (-1)^K \mu(d) \binom{\nu(d)}{K}.$$

Now let $\varphi_k(n) = \#\{m\colon 1 \leqslant m \leqslant n \text{ and } \nu((m,n)) < k\}$.

THEOREM 3.8. *For n square free, $\sum_{d\mid n} \mu_k(d)/d = \varphi_k(n)/n$.*

Proof.

$$n \sum_{d\mid n} \mu_k(d)/d = \sum_{d\mid n} \sum_{\{m:d\mid m,1\leqslant m\leqslant n\}} \mu_k(d) = \sum_{m=1}^{n} \sum_{d\mid(m,n)} \mu_k(d)$$

$$= \sum_{m=1}^{n} \begin{cases} 1 & \text{if } \nu((n,m)) < k \\ 0 & \text{if } \nu((n,m)) \geqslant k \end{cases} = \varphi_k(n).$$

4. A COMBINATORIAL SIEVE

Let us now recall that we are concerned with estimating $\#\mathscr{N}$ when $\mathscr{N} \subseteq (y, y + x]$ is sifted (k, z). Again let $\pi = \prod_{p<z} p$. For $1 \leqslant n \leqslant x$, $1 \leqslant \alpha \leqslant \pi$ let $\sigma_\alpha(n) = (n - \alpha, \pi)$. Then $\mathscr{N}$ is sifted (k, z) if and only if there exists α such that $\nu(\sigma_\alpha(n)) < k$ for $n \in \mathscr{N}$.

Accordingly fix x, z, k, and α, and let $\mathcal{N} = \{n: 1 \leqslant n \leqslant x$ and $\nu(\sigma_\alpha(n)) < k\}$. Note that $\#\{n: 1 \leqslant n \leqslant x,\ d \mid \sigma_\alpha(n)\} = x/d + \theta_d$ where $|\theta_d| < 1$ for all d. Now

$$\begin{aligned}\#\mathcal{N} &= \sum_{\substack{n=1\\ \nu(\sigma_\alpha(n)<k}}^{x} 1 = \sum_{n\leqslant x} \sum_{d\mid\sigma_\alpha(n)} \mu_k(d)\\ &= \sum_{d\mid\pi} \sum_{\{n:d\mid\sigma_\alpha(n)\}} \mu_k(d) = \sum_{d\mid\pi} \mu_k(d) \sum_{\{n:d\mid\sigma_\alpha(n)\}} 1\\ &= x\sum_{d\mid\pi} \mu_k(d)/d + \sum_{d\mid\pi} \theta_d = x\varphi_k(\pi)/\pi + \sum_{d\mid\pi} \theta_d M_k(d).\end{aligned}$$

If $2^{\pi(z)} > x$ we cannot dismiss the second term as "error." Thus we truncate by considering only d for which $\nu(d) < m - 1$, choosing m so that $k + m$ is even. Then

$$\sum_{n\leqslant x} \sum_{\substack{d\mid\sigma_\alpha(n)\\ \nu(d)<m-1}} \mu_k(d) \leqslant \#\mathcal{N} \leqslant \sum_{n\leqslant x} \sum_{\substack{d\mid\sigma_\alpha(n)\\ \nu(d)<m}} \mu_k(d),$$

by Theorem 3.6. From here it is clear how one should proceed to get concrete bounds.

We give details only for the case $k = 2$. Let

$$\sigma^{(m)}(d) = \sum_{\substack{\delta\mid d\\ \nu(\delta)<m}} \mu(\delta),$$

$s_m(d) = \mu(d)\,\sigma^{(m)}(d)$, and $\lambda_2(d) = \mu(d)\,\mu_2(d)$. Note that $\sigma^{(m)}(d) = (-1)^{m-1}\binom{\nu(d)-1}{m-1} = \lambda_m(d)$. Let $\chi_m(d) = 1$ if $\nu(d) < m$, else 0. By Möbius inversion, $\chi_m(d) = \sum_{\delta\mid d} s_m(\delta)$. The main term of $\#\mathcal{N}$ is $x\sum_{d\mid\pi} \mu_2(d)/d$.

Now

$$\begin{aligned}\sum_{d\mid\pi,\nu(d)<m} \mu_2(d)/d &= \sum_{d\mid\pi} \frac{1}{d} \sum_{\delta\mid d} \mu(d)\, s_m(\delta)\, \lambda_2(d)\\ &= \sum_{d\mid\pi} \frac{1}{d} \sum_{\delta\mid d} \mu\,\frac{d}{\delta}\, \sigma^{(m)}(\delta)\cdot\left(\lambda_2\left(\frac{d}{\delta}\right) - \nu(\delta)\right)\\ &= \sum_{\delta\mid\pi} \frac{\sigma^{(m)}(\delta)}{\delta} \sum_{t\mid\pi/\delta} \frac{\mu(t)}{t}\,(\lambda_2(t) - \nu(\delta)).\end{aligned}$$

Now let $W(z) = \prod_{p<z}(1 - (1/p))$, $T(z) = 1 \times \sum_{p<z}(1/(p-1))$, $T_\delta = \sum_{p\mid\delta}(1/(p-1))$, and $L_\delta \sum_{p\mid\delta}(p/(p-1))$. Note that $\varphi_2(n) = \varphi(n)(1 + T_n)$. Our chain of equalities for the main term now continues:

$$= -\sum_{\delta|\pi} \frac{\sigma^{(m)}(\delta)\,\nu(\delta)}{\delta} \sum_{t|\pi/\delta} \frac{\mu(t)}{t} + \sum_{\delta|\pi} \frac{\sigma^{(m)}(\delta)}{\delta} \sum_{t|\pi/\delta} \frac{\mu_2(t)}{t}$$

$$= -\sum_{\delta|\pi} \frac{\sigma^{(m)}(\delta)\,\nu(\delta)\,\varphi(\pi/\delta)}{\pi} + \sum_{\delta|\pi} \frac{\sigma^{(m)}(\delta)\,\varphi_2(\pi/\delta)}{\pi}$$

$$= -W(z) \sum_{\delta|\pi} \frac{\sigma^{(m)}(\delta)\,\nu(\delta)}{\varphi(\delta)} + W(z)\,T(z) \sum_{\delta|\pi} \frac{\sigma^{(m)}(\delta)}{\varphi(\delta)} - W(z) \sum_{\delta|\pi} \frac{\sigma^{(m)}(\delta)\,T_\delta}{\varphi(\delta)}$$

$$= W(z)\,T(z) \sum_{\delta|\pi} \frac{\sigma^{(m)}(\delta)}{\varphi(\delta)} - W(z) \sum_{\delta|\pi} \frac{\sigma^{(m)}(\delta)\,L_\delta}{\varphi(\delta)}.$$

Now following Halberstam and Richert [2, pp. 48–50] this is (when multiplied by x)

$$= xW(z)\,T(z)\{1 + O(e^{-(\log x)^{1/2}})\} \qquad \text{for} \quad \log z \leqslant (\log x)^{1/2}.$$

The error term $\sum_{d|\pi} \theta_d M_k(d)$ is, for $z \geqslant 14$, $O(X^{1/2})$. So we have

THEOREM 4.1. *If $z \leqslant 3 \log z \leqslant (\log x)^{1/2}$, $\alpha \in Z$, and $\mathcal{N} = \{n\colon 1 \leqslant n \leqslant x$ and $\nu(n - \alpha, \pi) < 2\}$ then*

$$\#\mathcal{N} = xW(z)\,T(z)(1 + O(e^{-(\log x)^{1/2}})) + O(x^{1/2}).$$

COROLLARY. *$\#\mathcal{N} \approx e^{-\gamma} x \log\log z/\log z$ under the above hypotheses.*

Proof. Recall the definitions of W and T.

5. A SELBERG-TYPE SIEVE

For the sake of simplicity we confine ourselves again to the case $k = 2$. Fix x and z, and let $\pi = \prod_{p<z} p$. Let $\mathcal{D}$ be a divisor-closed subset of the divisors of π. Our objective is to choose, in place of

$$s_0(n) = \sum_{\substack{d|\sigma(n)\\ d|\pi}} \mu_2(d),$$

a function

$$s(n) = \Bigl\{\sum_{\substack{d|\sigma(n)\\ d\in\mathcal{D}}} \Lambda_d\Bigr\}^2$$

so that $s(n) \geqslant s_0(n)$ but gives a smaller error term. Following Selberg's method we would like to minimize

$$H(\Lambda) = \sum_{\delta\in\mathcal{D}} \varphi(\delta)\, y_\delta^2 \qquad \text{where} \quad y_\delta = \sum_{\substack{d\in\mathcal{D}\\ \delta|d}} \frac{\Lambda_d}{d},$$

under the constraint $\Lambda_1 = 1$, $\Lambda_p = 0$. With these constraints the exact minimum of $H(\Lambda)$ occurs when the vector

$$(\Lambda_d/d), \ \underset{\substack{d\in\mathscr{D}\\ \nu(d)\geq 2}}{\text{for}}$$

satisfies a certain matrix equation. Let $M = M(\mathscr{D})$ be the matrix with rows and columns corresponding to elements d of $\mathscr{D}$ with $\nu(d) \geq 2$, and with the entry at the row and column corresponding to d_1, d_2, respectively, equal to gcd $(d_1, d_2) = (d_1, d_2)$. Thus, for instance, if $\mathscr{D} = \{1, 2, 3, 5, 6, 7, 10, 14, 15\}$,

$$M = \begin{array}{c} \\ 6 \\ 10 \\ 14 \\ 15 \end{array}\begin{array}{c} \begin{array}{cccc} 6 & 10 & 14 & 15 \end{array} \\ \begin{pmatrix} 6 & 2 & 2 & 3 \\ 2 & 10 & 2 & 5 \\ 2 & 2 & 14 & 1 \\ 3 & 5 & 1 & 15 \end{pmatrix} \end{array}.$$

Then (Λ_d/d) minimizes $H(\Lambda)$ under the above constraints if and only if $M(\Lambda_d/d) = (-1) \cdot$ (vector identity). Since the exact minimum of $H(\Lambda)$ appears to be quite complex, we shall choose Λ_d by educated guesswork. Let

$$Q_d = \sum_{\substack{t\in\mathscr{D}\\ d|t}} \frac{1}{\varphi(t)}, \qquad Q = Q_1 = \sum_{t\in\mathscr{D}} \frac{1}{\varphi(t)}.$$

Now take $\Lambda_d = d\mu_2(d)\, Q_d/Q$ (in the strict Selberg sieve one has here μ, not μ_2).

If $\mathscr{D}$ consists of all divisors of π then this choice of Λ_d gives $s(n) = s_0(n)$. The error, then, results from choosing a smaller $\mathscr{D}$ in an attempt to control the other, "roundoff" error.

For $t \in \mathscr{D}$ let

$$Q(\mathscr{D}/t) = \sum_{s\in(\mathscr{D}/t)} \frac{1}{\varphi(s)} \qquad \text{where} \quad \mathscr{D}/t = \{d/t\colon t \mid d \text{ and } d \in \mathscr{D}\}.$$

Then

$$y_\delta = \frac{\mu(\delta)}{Q\varphi(\delta)} \sum_{d'\in\mathscr{D}/\delta} \mu(d')(1 - \nu(\delta) - \nu(\delta')) \cdot Q(\mathscr{D}/d'\delta)/\varphi(d').$$

Note that

$$\sum_{t\in\mathscr{S}} \frac{\mu(t)\, Q(\mathscr{S}/t)}{\varphi(t)} = 1$$

for any divisor-closed $\mathscr{S}$. Thus

$$y_\delta = \frac{\mu(\delta)(1-\nu(\delta))}{Q\varphi(\delta)} + \frac{\mu(\delta)}{Q\varphi(\delta)} \sum_{p\in\mathscr{D}/\delta} \frac{1}{p-1} \sum_{d''\in\mathscr{D}/\delta_p} \mu(d'')\, Q(\mathscr{D}/d''p\delta)/\varphi(d'')$$

$$= \frac{\mu(\delta)}{Q\varphi(\delta)}\left(1-\nu(\delta)+\sum_{\substack{p\in\mathscr{D}\\ p\nmid\delta}}\frac{1}{p-1}\right).$$

With

$$T = \left(1+\sum_{p\in\mathscr{D}}\frac{1}{p-1}\right),$$

and L_δ as before,

$$y_\delta = \frac{\mu(\delta)}{Q\varphi(\delta)}(T-L_\delta),$$

and

$$H(\Lambda) = \sum_{\delta\in\mathscr{D}} \varphi(\delta)\, y_\delta{}^2 = (1/Q^2)\sum_{\delta\in\mathscr{D}} (T-L_\delta)^2/\varphi(\delta).$$

If we now take $\mathscr{D}$ of the form $\{d \leqslant z: \mu^2(d) = 1\}$, and $z^2 \leqslant x/\log^2 x$ this last can be estimated. The analysis here was suggested by Selberg.

THEOREM 5.1. *With Λ_d as above,*

$$H(\Lambda) = \frac{\log\log z}{\log z} + O\left(\frac{1}{\log z}\right).$$

Proof. Let ξ be a complex variable, and note that

$$H(\Lambda) = Q^{-2}\sum_{\mathscr{D}}\frac{(T-L_\delta)^2}{\varphi(\delta)}$$

$$= Q^{-2}\frac{\partial^2}{\partial\xi^2}\left\{e^{\xi T}\sum_{\mathscr{D}} e^{-\xi L_\delta/\varphi(\delta)}\right\}.$$

Let $F_1(z) = \sum^*_{\delta\leqslant z}(T-L_\delta)^2/\varphi(\delta)$, where $\sum^*$ denotes a sum restricted to square-free numbers. Let $F(\xi, z) = \sum^*_{\delta\leqslant z} e^{-\xi L_\delta/\varphi(\delta)}$, and $f(\xi, s) = \sum^{*\infty}_{\delta=1} e^{-\xi L_\delta/\varphi(\delta)\delta^s} = \prod_p (1+e^{-\xi_p/\varphi(p)p^s})$. Let $g(\xi, s)$ satisfy $f(\xi, s) = (\zeta(1+s))^{e^{-\xi}} g(\xi, s)$. Then $g(\xi, s)$ is uniformly convergent for ξ in a neighborhood of 0, and $\operatorname{Re} s \geqslant -\frac{2}{5}$.

We now state some familiar lemmas of analysis. In some cases the proofs are omitted.

LEMMA 5.1. *If $G(x)$ is a function on $(0, \infty)$ vanishing on $(0, 1)$ and*

$\operatorname{Var} G(0, x) = O(x^\epsilon)$ for all $\epsilon > 0$, and if $g(s) = \int_0^\infty x^{-s}\, dG(x)$ for $\operatorname{Re}(s) > 0$ *then for $c >$, T real,*

$$G(x) = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} g(s)\, x^s\, ds/s + O_c(\Delta(x)) + O_c\left(\frac{1}{T}(x \log x\, \Delta(x) + x^c)\right),$$

with

$$\Delta(x) = \operatorname{Var} G(x-1, x+1).$$

LEMMA 5.2. *If C is the contour going from $-\infty$ to $-\epsilon$, then counter clockwise around a circle of radius ϵ back to $-\epsilon$, and then from $-\epsilon$ to $-\infty$, then*

$$\frac{1}{2\pi i}\int_C x^s s^{-\alpha}\, ds = (\log x)^{\alpha-1}/\Gamma(\alpha) \qquad \text{for} \quad \operatorname{Re}(\alpha) \geqslant 1.$$

LEMMA 5.3. *The error in truncating the integral above by removing the segments from $-\delta$ to $-\infty$ is $O_\delta(x^{-\delta}/\log x)$.*

LEMMA 5.4. *Let $h(\xi, s) = (s\zeta(1+s))^{e^{-\xi}} f(\xi, s)$. Then $h(\xi, s) = h(\xi, 0) + sh_s(\xi, 0) + O(s^2)$ uniformly for ξ in a neighborhood of zero. (Here $h_s = \partial h/\partial s$).*

LEMMA 5.6. *Uniformly in a ξ-neighborhood of zero,*

$$F(\xi, z) = (\log z)^{e^{-\xi}}\left(\frac{h(\xi, 0)}{\Gamma(1 + e^{-\xi})}\right) + O(\log z)^{1/2}.$$

Proof.

$$F(\xi, z) = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{z^s}{s} f(\xi, s)\, ds + O(1)$$

from Lemma 5.1 on taking $T = z^c$ and noting that

$$\Delta(z) = O\left(\frac{\log\log z}{z}\right).$$

Now let C_1 be the contour from $c - iT$ to $c + iT$, then to $-\delta - iT$, then to $-\delta$, then along C backwards and back out to $-\delta$, then to $-\delta - iT$, and closing at $c - iT$. Break C_1 up into several pieces: I_1 from $c - iT$ to $c + iT$, I_2 the portion along C, and I_3, the rest. Then

$$\int_{C_1} \frac{z^s}{s} f(\xi, s)\, ds = 0$$

by Cauchy's theorem.

We want to estimate

$$F(\xi, z) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{z^s}{s} f(\xi, s)\, ds.$$

The error in truncating at $\pm iT$ has been accounted for by Lemma 5.1. The backstretch component $\int_{I_3} |(z^s/s) f(\xi, s)|$ is $O(\log z)^{-1/2} \cdot | \int_{I_2} - \int_C | = O(\log z)^{-1/2}$. So it remains to evaluate

$$\int_C \frac{z^s}{s} f(\xi, s)\, ds = \int_C \frac{z^s}{s^{(1+e^{-\xi})}} (h(\xi, 0) + sh_s(\xi, 0) + O(s^2)).$$

The term due to $h(\xi, 0)$ is the ain term in Lemma 5.6, from Lemma 5.2.

For the other terms, let us take the radius ϵ to be $1/\log z$. Then

$$\int_C \frac{z^s}{s^{(1+e^{-\xi})}} sh_s(\xi, 0)\, ds = \frac{h_s(\xi, 0)(\log z)^{(e^{-\xi}-1)}}{\Gamma(e^{-\xi})} = O(\log z)^{1/2}$$

uniformly in a ξ-neighborhood of 0. The δ-truncation errors are $O(z^{-\delta})$, less than $(\log z)^{1/2}$.

Finally

$$\int_{I_3} O(s^2) \cdot \left| \frac{z^\epsilon \text{ darc length}(s)}{s^{(1+e^{-\xi})}} \right| = O(\log z^{(1/2)}),$$

with the radius $\epsilon = 1/\log z$. This proves Lemma 5.6.

Now let

$$G(\xi, z) = \frac{h(\xi, 0))\log z)^{(e^{-\xi})}}{\Gamma(1 + e^{-\xi})}.$$

Then with $T = T(z)$,

$$\begin{aligned} F_1(z)\ & T^2F(0, z) - 2TF_\xi(0, z) + F_{\xi\xi}(0, z) \\ &= T^2G(0, z) - 2TG_\xi(0, z) + G_{\xi\xi}(0, z) + O(\log z)^{1/2} \\ &= \log \log z \log z + O(\log z), \end{aligned}$$

since

$$T = \log \log z + k + O\left(\frac{1}{\log z}\right),$$

where k is a known constant. The 0-estimate of Lemma 5.6 can be differentiated because it is uniform in a ξ-neighborhood of zero and Cauchy's principle applies. This completes the proof of Theorem 5.1.

We still have the roundoff error to consider. Let $E = \sum_{d_1 \in \mathscr{D}} \sum_{d_2 \in \mathscr{D}} | \Lambda_{d_1} \Lambda_{d_2} R_{[d_1, d_2]} |$. Then by Selberg's sieve theorem $\# \mathscr{N} \leqslant xH(\Lambda) + E$. Now in our case $| R_d | \leqslant 1$, $| \Lambda_d | \leqslant d(1 + \nu(d)) Q_d/Q \leqslant d(1 + \nu(d))/\varphi(d)$, so

$$E = O\left(\sum_{\mathscr{D}} \frac{d(1 + \nu(d))}{\varphi(d)}\right)^2 = O\left(z^2 \log^2 z \left(\sum_{d \leqslant z}^* 1/\varphi(d)\right)^2\right) = O(z^2 \log^4 z),$$

as $\nu(d) \ll \log d$ and $\sum_{d \leqslant z}^* 1/\varphi(d) = \log z + O(1)$. Thus we have proved

THEOREM 5.2. *If* $\{1, 2, \ldots, [x]\}$ *is sifted* $(2, z)$ *to form* $\mathcal{N}$, *then* $\#\mathcal{N} \leqslant x \log\log z/\log z + O(x/\log z) + O(z^2 \log^4 z)$.

REFERENCES

1. A. BAKER, "Transcendental Number Theory," Theorem 3.1, Cambridge Univ. Press, London, 1974.
2. H. HALBERSTAM AND H.-E. RICHERT, "Sieve Methods," pp. 48–50, Academic Press, New York/London, 1974.
3. H. MONTGOMERY AND R. VAUGHAN, The large sieve, *Mathematika* **20** (1973), 119–134.